Personality Disorders in Croatia's Cyber Political Networks – A Topic Map Study

Gordan Ponjavić
TiAktiv
Udruga za demokratizaciju politike
10000 Zagreb, Croatia
gponjavic@gmail.com

Markus Schatten
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
markus.schatten@foi.hr

Miroslav Bača
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
miroslav.baca@foi.hr

Frane Jakelić
Undergraduate student
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
frane.jakelic@foi.hr

Abstract. *Recent research has shown a correlation between personality disorders (PDs) and the position of a person in a social network. Following this research we conducted a social network analysis on one of the greatest Croatian cyber-political blogging sites to identify individuals with possible PDs. These individuals' blogs were then analyzed in order to identify topics which are associated with various types of disorders.*

Keywords. Croatia, cyber politics, social network analysis, personality disorder

1 Introduction

Clifton et al. suggested that there is a correlation between a persons position in a social network and various kinds of PDs [1]. This amazing findings motivated us to analyze further and to see if one can identify certain political topics associated with various PDs.

In this paper we analyzed one of the major Croatian political blogging sites. The site was harvested from its start-up until November 11th 2009 for its social network (a connection was established on each comment to some post) as well as for the content of posts and comments.

The network was then analyzed for the various measures indicated in [1] to be associated with various PDs. Identified actors' posts and comments where then analyzed further using text mining techniques explained later. Results from this concept analysis were then summarized into topic maps for various network measures.

In the following we will first describe shortly the various PDs that were analyzed in section 2. In section 3 we give a brief outline of social network analysis with accent on the network metrics that were used for this study. Section 4 explains the Y!Q contextual search system [3] as well as topic maps created with the ICTA tool [2]. In section 5 we present the methodology and the results of our investigation. Section 6 provides a discussion and concluding remarks.

2 Personality Disorders

Personality disorders (PDs) are a class of personality types and behaviors (patterns) that the American Psychiatric Association (APA) defines as "an enduring pattern of inner experience and behavior that deviates markedly from the expectations of the culture of the individual who exhibits it"[1]. The following PDs were analyzed in this study:

Paranoid Pervasive suspiciousness that others are trying to harm or exploit him or her

Schizoid Emotional coldness and social isolation

Schizotypal Eccentric behavior, cognitive and perceptual abnormalities, social withdrawal

Antisocial Violation of laws, morality, and the rights of others

Histrionic Attention seeking, over-exaggerated expression of emotion

Narcissistic Grandiosity, feelings of entitlement, lack of empathy for others

Avoidant Extreme shyness, social inhibition, fear of evaluation

Dependent Need for reassurance and care-taking by others

Obsessive-Compulsive PD Perfectionism and need for control

3 Social Network Analysis

Social network analysis is concerned with understanding the connections among social entities as well as with the implications of such linkages [5, pp. 17–20.]. A more formal approach to defining social networks is graph theory.[1]

Definition A *graph* $\mathcal{G}$ is the pair $(\mathcal{N}, \mathcal{E})$ whereby $\mathcal{N}$ represents the set of *verticles* or *nodes*, and $E \subseteq N \times N$ the set of *edges* connecting pairs from $\mathcal{N}$.

A graph can be represented with the so called adjacency matrix.

Definition Let $\mathcal{G}$ be a graph defined with the set of nodes $\{n_1, n_2, ..., n_m\}$ and edges $\{e_1, e_2, ..., e_l\}$. For every i, j $(1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant m)$ we define

$$a_{ij} = \begin{cases} 1, & \text{if there is an edge between nodes} \\ & n_i \text{ and } n_j \\ 0, & \text{otherwise} \end{cases}$$

Matrix $A = [a_{ij}]$ is then the adjacency matrix of graph $\mathcal{G}$. The matrix i symmetric since if there is an edge between nodes n_i and n_j then clearly there is also an edge between n_j and n_i. Thus $A = [a_{ij}] = [a_{ji}]$.

The notion of directed graphs and directed multi-graph are of special importance to our study.

Definition A *directed graph* or *digraph* $\mathcal{G}$ is the pair $(\mathcal{N}, \mathcal{A})$, whereby $\mathcal{N}$ represents the set of *nodes*, and $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ the set of ordered pairs of elements from $\mathcal{N}$ that represent the set of graph *arcs*.

Definition A *directed multi-graph* or *multi-digraph* $\mathcal{G}$ is the pair $(\mathcal{N}, \mathcal{A})$, whereby $\mathcal{N}$ represents the set of *nodes*, and $\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$ the multi-set of ordered pairs of elements from $\mathcal{N}$ that represent the set of graph *arcs*.

A social network can be represented as a graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ where $\mathcal{N}$ denotes the set of actors, and $\mathcal{A}$ denotes the set of relations between them [4]. If the relations are directed (e.g. support, influence, message sending etc.) we can conceptualize a social network as a directed graph. If there is possibly more than one connection between two nodes (as is the case in message sending) then the social network can be represented as a multi digraph.

One of the main applications of graph theory to social network analysis is the identification of "most important" actors inside a social network. There a lots of different methods and algorithms that allow us to calculate the importance, prominence, degree, closeness, betweenness, information, differential status or rank of an actor.[2]. Herein we would like to outline four such measures which are in-degree (the number of incoming relations of some actor), out-degree (the number of outgoing connections of some actor), in-degree-out-degree (the difference between the previous two measures) and betweenness-centrality (the probability that a given node lies on a shortest paths connecting two other nodes).

Clifton et al. showed that there is correlation between PDs and the position of a person in a social network as shown in table 1 [1].

[1] There are off course other approaches like socio-metrics.

[2] See [5] for an in depth discussion of such metrics.

Table 1: Personality disorders with correlated social network metrics [1] whereby + indicates positive correlation, − negative correlation and empty cells indicate no significant correlation

PD	Centrality	In-degree	Out-degree	Indegree-Out-degree
Paranoid			+	−
Schizoid	−	−	−	−
Schizotypal		−	−	
Antisocial	+		+	−
Histrionic	+		+	
Narcissistic	+		+	−
Avoidant	−	−	−	
Dependent			−	
OCPD			+	−

4 Contextual Search and Topic Maps

In this study contextual search (namely the Y!Q large scale contextual search system [3]) was used to identify topic patterns inside the communication between actors of the blogging network. Contextual search augments the search context of user's queries in order to capture and find better information. The Y!Q system comprises of three major components:

- **Content analysis** - performs a content analysis of a given context by exposing a semantic network.
- **Query planning and rewriting framework** - tries to identify the best places to search.
- **Contextual ranking** - re-ranks the obtained results.

The greatest benefit from contextual search is that it helps to disambiguate queries [3].

After identifying most relevant topics using Y!Q the ICTA (Internet Community Text Analyzer) [2] tool was used to create topic maps. Topic maps show the development of topics through time. The larger the number of times a certain topic was used in some text, the surface on the topic map.

5 Methodology & Results

In order to harvest the social network a little Python[3] script was implemented that downloaded all messages (a total of 183,016 messages) from the blogging site into a PostgreSQL[4] database. The acquired data was later on analyzed first by using the NetworkX[5] Python module for large scale network analysis to calculate nodes' centrality, in-degree, out-degree and in-degree-out-degree. According to the correlations presented in table 1, the top and bottom 10 % of actors were included into further analysis in order to identify specific topics. The identified actors' messages were then analyzed using Y!Q and ICTA to create topic maps presented on figures 1, 2, 3, 4, 5 and 6, showing topic maps of users with low centrality, high centrality, low in-degree, low out-degree, high out-degree and low in-degree-out-degree respectively.

Messages of users with low centrality (a total of 2,372 messages), low in-degree (a total of 2,375), and low out-degree (a total of 387 messages) were analyzed in their entirety. Due to the fact that the other message sets were relatively big (high centrality - 148,646 messages; high out-degree 170,793 messages; low in-degree-out-degree 180,907 messages) and due to our limited resources some message sets were only partially analyzed using random samples as indicated on figures 2, 5 and 6.

[3] http://www.python.org
[4] http://www.postgresql.org
[5] http://networkx.lanl.gov/

6 Discussion & Conclusion

As one can see on figure 1 bloggers with low centrality (which correlates with schizoid and avoidant PDs) are often concerned with law enforcement (cro. prava - rights; cro. pravo - right/law; cro. sabor - parliament; cro. hrvatski sabor - Croatian parliament; cro. sudu, suda - court; cro. suca, sudaca - judge, judges; cro. ustav - constitution; cro. država - state). Such bloggers seem to talk more about the state's internal affairs.

On the other hand bloggers with high centrality (which correlates with antisocial, histrionic and narcissistic PDs) as seen on figure 2, are more concerned with external affairs (topics like NATO, Iran, Bush, and a lot of topics dealing with Bosnia & Herzegovina are self explanatory).

The map of bloggers' topics with low in-degree (similar to low centrality correlating with schizoid, schizotypal, and avoidant) is shown on figure 3. Again such actors seem to be concerned with law enforcement (prava, pravo, suda, sudu, sude, sudaca, sabora, hrvatski sabor, ustav, cro. udruga - association/society/organization) but here we also find few patterns concerning external affairs (Kosovo, NATO).

Bloggers with low out-degree (correlated with schizoid, schizotypal, avoidant, and dependent) seem to be talking a lot about controversial topics (Hitler; Homo; NATO; Bruxelles; Thomson - Croatian nationalist-oriented pop singer; cro. nacionalizam - nationalism) as shown on figure 4. Few topics also concern time (cro. danas - today; cro. dana - days; cro. vrijeme - time) and state politics/law enforcement (cro. vlada - government; Sanader - Croatia's ex-prime minister; cro. država - state; cro. grad Zagreb - the city of Zagreb).

Figure 5 shows the topic map of bloggers with high out-degree (correlated with paranoid, antisocial, histrionic, narcissistic and obsessive-compulsive PDs). In this topic map a pattern is hard to find which could be due to inadequate sampling. Groups of topics include external affairs (Giuliani; Ground Zero; cro. Izrael - Israel; Bechtel; NATO), nationalistic topics & religion (Thompson; cro. vijenac - corolla/wreath; cro. Boga - God), state politics (HSU acronym for cro. Hrvatska Stranka Umirovljenika - Croatian Pensioner Party; referendum)

The topic map of low in-degree-out-degree (correlating with paranoid, schizoid, antisocial, narcissistic and obsessive-compulsive PDs) is shown on figure 6. Such bloggers talk about economy and money/income affairs (Keynes; cro. posao - job/-work; cro. kapitalizam - capitalism; risk management; cro. stan - apartment).

From this analysis we can conclude that there is a noticeable difference between political bloggers with low centrality (talking mostly about internal affairs and correlated with introverted PDs like schizoid and avoidant), and political bloggers with high centrality (talking mostly about external affairs and correlated with extroverted PDs like antisocial, histrionic and narcissistic). There also seems to be a pattern that bloggers which position in their social network indicates correlation to schizoid, schizotypal, and avoidant PDs, seem to like to talk about law enforcement (the pattern is perceptible where bloggers have low centrality, low in-degree and low out-degree). The other correlations indicated in [1] did not seem to have counterparts in topic maps.

References

[1] Clifton, A., Turkheimer, E., and Oltmanns, T. F. Personality disorder in social networks: Network position as a marker of interpersonal dysfunction. *Social Networks 31* (2009), 26–32.

[2] ICTA. Internet Community Text Analyzer. available at `http://textanalytics.net`, Accessed: 3^{rd} March 2010.

[3] Kraft, R., Maghoul, F., and Chang, C. C. Y!Q: contextual search at the point of inspiration. In *CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management* (New York, NY, USA, 2005), ACM, pp. 816–823.

[4] Mika, P. *Social Networks and the Semantic Web*. Springer, New York, 2007.

[5] Wasserman, S., and Faust, K. *Social Network Analysis ; Methods and Applications*. Structural analysis in the social sciences. Cambridge University Press, 1994.

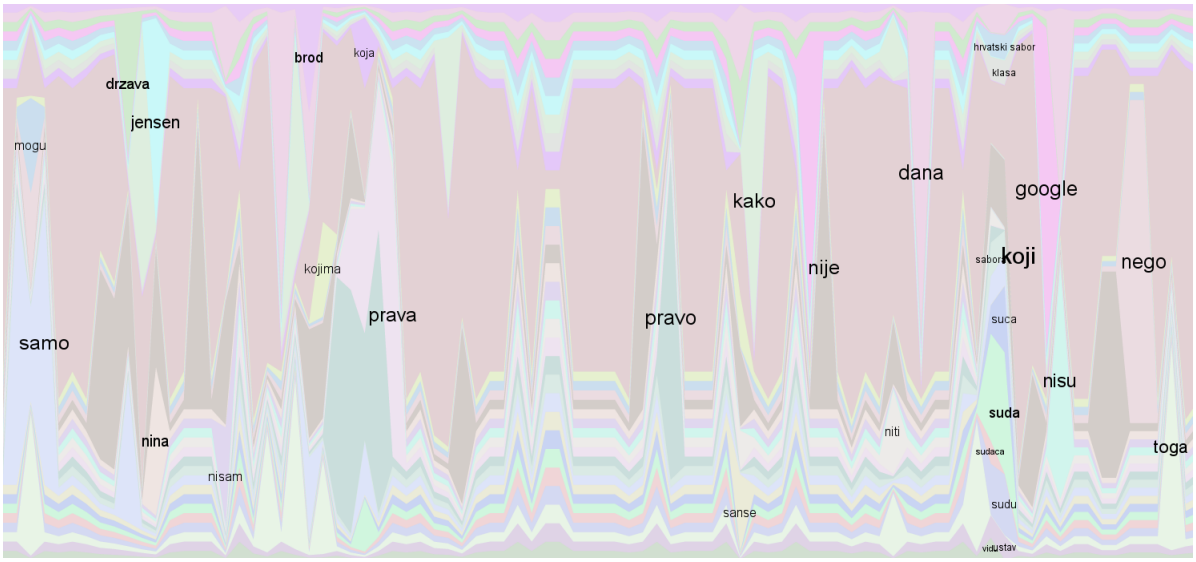

Figure 1: Low centrality topic map (100% of processed messages)

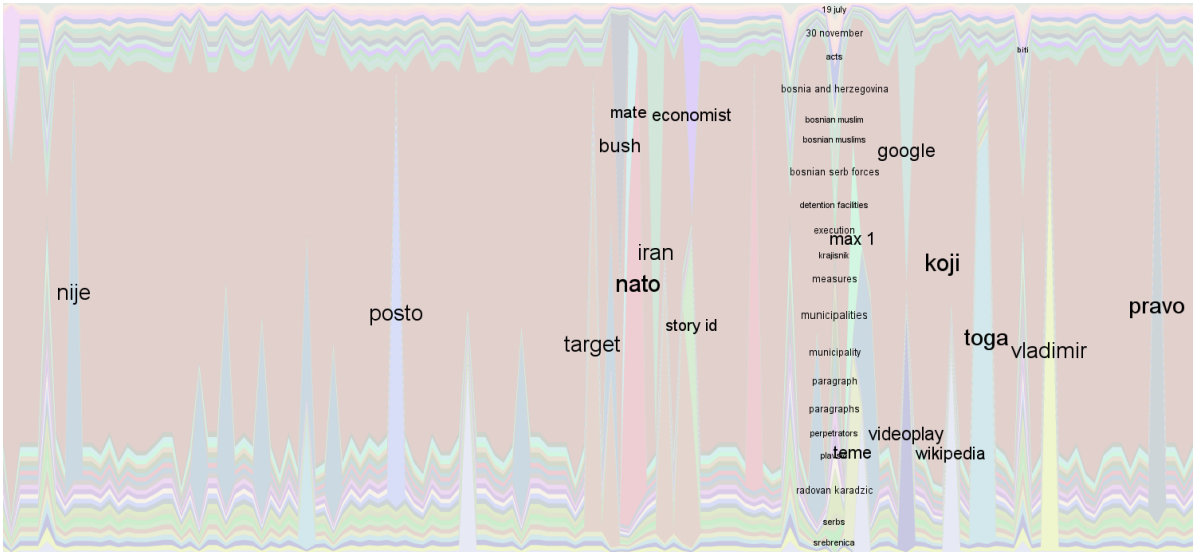

Figure 2: High centrality topic map (50% of processed messages)

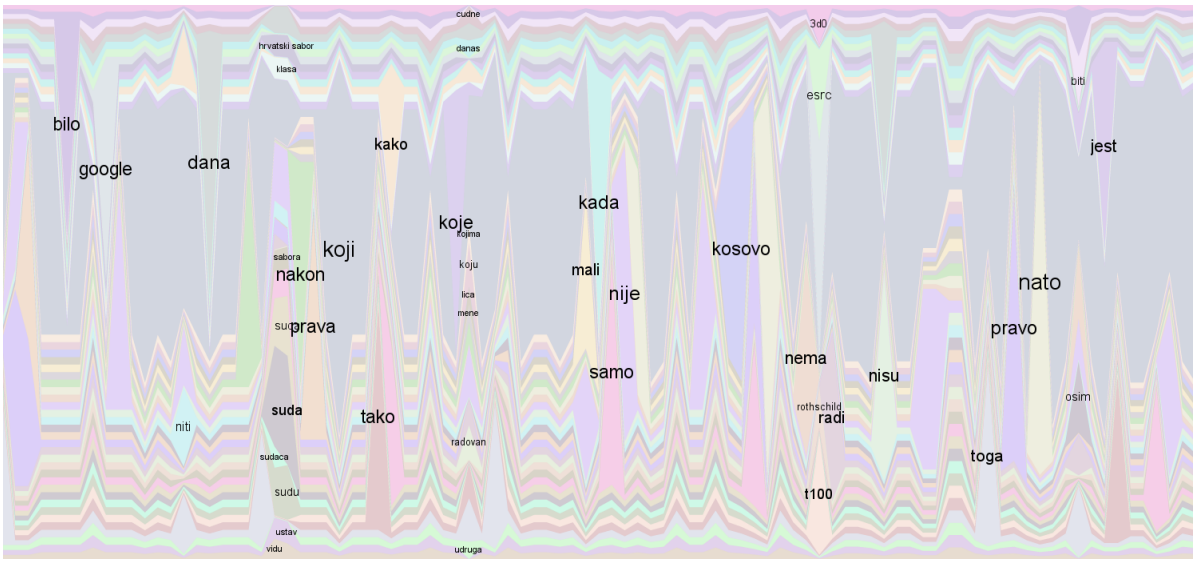

Figure 3: Low indegree topic map (100% of processed messages)

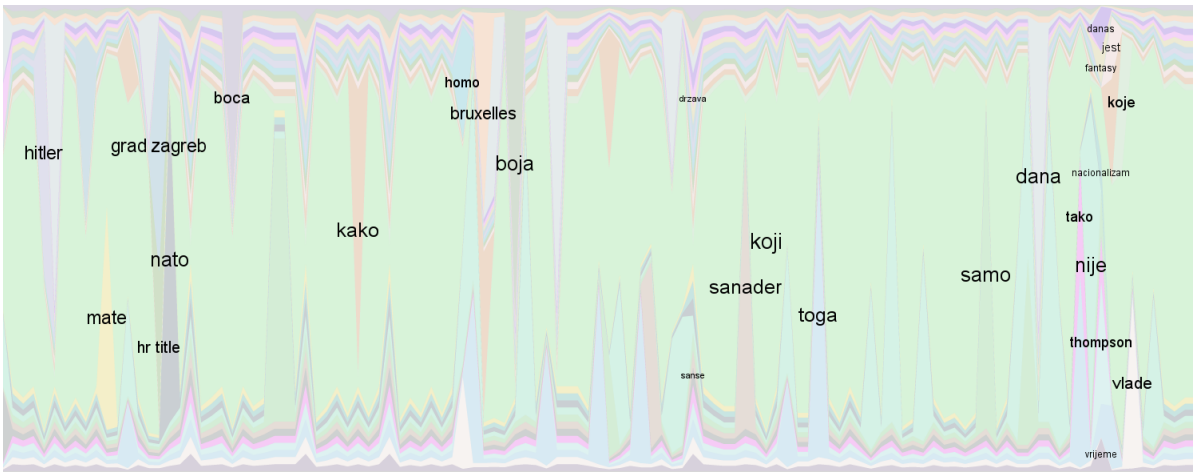

Figure 4: Low outdegree topic map (100% of processed messages)

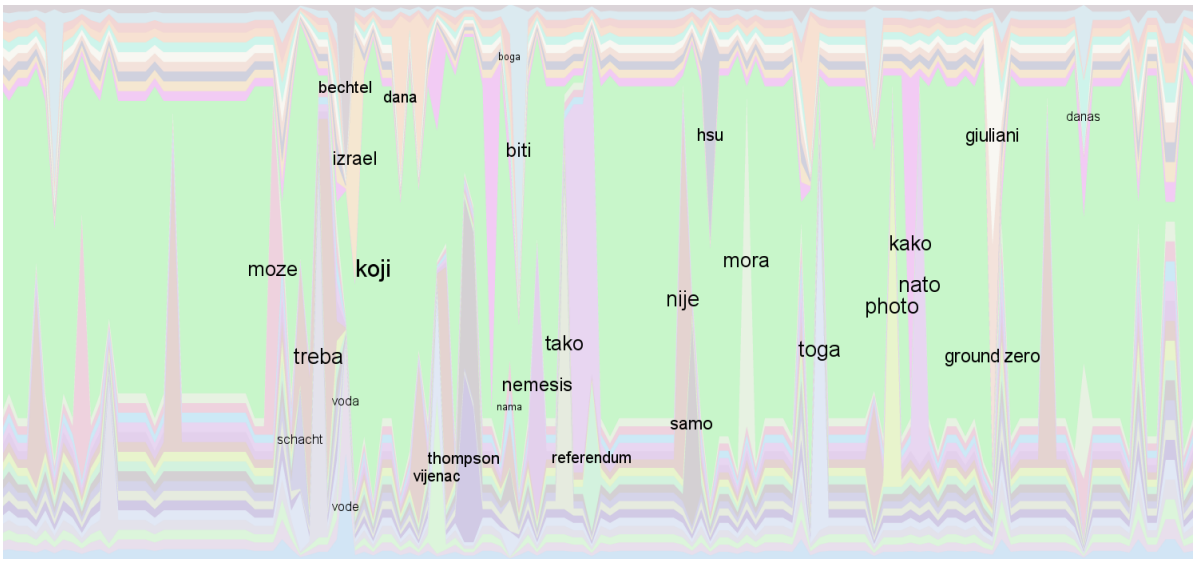

Figure 5: High outdegree topic map (7% of processed messages)

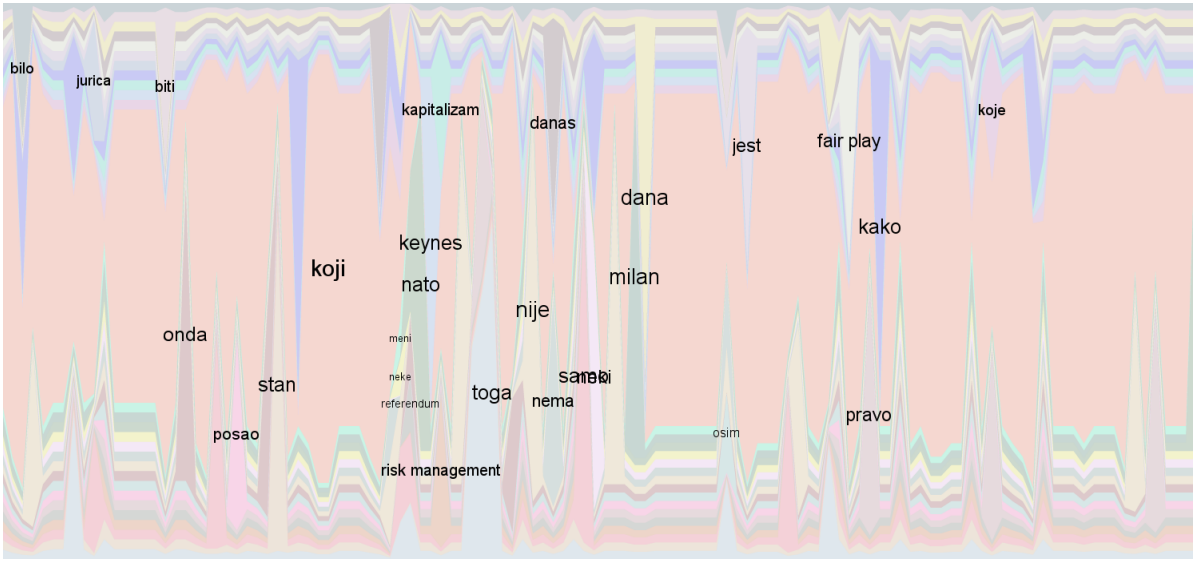

Figure 6: Low indegree - outdegree topic map (5% of processed messages)